ELSEVIER Stochastic Processes and their Applications 60 (1995) 171–190

stochastic processes and their applications

Comparing Fleming–Viot and Dawson–Watanabe processes

S.N. Ethier [a,1], Stephen M. Krone [b,*]

[a] *Department of Mathematics, University of Utah, Salt Lake City, UT 84112, USA*
[b] *Department of Mathematics, University of Wisconsin, Madison, WI 53706, USA*

Abstract

Fleming–Viot processes and Dawson–Watanabe processes are two classes of "superprocesses" that have received a great deal of attention in recent years. These processes have many properties in common. In this paper, we prove a result that helps to explain why this is so. It allows one to prove certain theorems for one class when they are true for the other. More specifically, we show that product moments of a Fleming–Viot process can be bounded above by the corresponding moments of the Dawson–Watanabe process with the same "underlying particle motion", and vice versa except for a multiplicative constant. As an application, we establish existence and continuity properties of local time for certain Fleming–Viot processes.

AMS Subject Classification: Primary: 60G57, 60J60; Secondary: 60J55, 60G17

Keywords: Fleming–Viot; Dawson–Watanabe; Superprocess; Measure-valued process; Moments; Local time

1. Introduction

Fleming–Viot processes are probability-measure-valued Markov processes that arise as diffusion approximations for many Markov chains in population genetics. These processes were introduced by Fleming and Viot (1979) and have been studied by Dawson and Hochberg (1982) and others. See Ethier and Kurtz (1993) for a recent survey. To define such a process, let E be a locally compact separable metric space (representing the set of alleles or "types" in genetics applications), and let $\mathscr{P}(E)$ denote the set of Borel probability measures on E with the topology of weak convergence. Let A be the generator of a Feller semigroup on $\hat{C}(E)$, the space of continuous functions on E vanishing at infinity. The Fleming–Viot process with mutation operator A (or the A-FV process) is a $\mathscr{P}(E)$-valued Markov process with a

* Corresponding author.

[1] Supported in part by NSF grants DMS–9102925 and DMS–9311984.

SSDI 0304-4149(95)00056-9

generator given by

$$(\mathscr{L}^{\mathrm{FV}}\varphi)(\mu) = \frac{1}{2}\int_E\int_E \mu(\mathrm{d}x)(\delta_x(\mathrm{d}y) - \mu(\mathrm{d}y))\frac{\delta^2\varphi(\mu)}{\delta\mu(x)\,\delta\mu(y)} + \int_E \mu(\mathrm{d}x)A\Big(\frac{\delta\varphi(\mu)}{\delta\mu(\cdot)}\Big)(x), \tag{1.1}$$

where $\delta\varphi(\mu)/\delta\mu(x) = \lim_{\varepsilon\to 0+}\varepsilon^{-1}\{\varphi(\mu+\varepsilon\delta_x) - \varphi(\mu)\}$ and $\delta_x \in \mathscr{P}(E)$ denotes the unit mass at $x \in E$. Its domain can be taken to be $\mathscr{D}(\mathscr{L}^{\mathrm{FV}}) = \{\varphi : \varphi(\mu) \equiv F(\langle\mu, f_1\rangle, \ldots, \langle\mu, f_k\rangle), F \in C^2(\mathbf{R}^k), f_1, \ldots, f_k \in \mathscr{D}(A), k \geqslant 1\}$, where $\langle\mu, f\rangle = \int_E f\,\mathrm{d}\mu$.

Dawson–Watanabe processes are finite-measure-valued Markov processes that arise as high-density limits of systems of branching and diffusing particles. They were introduced by Watanabe (1968) and Dawson (1975), and have since been studied by many others. One can consult Fitzsimmons (1988) for the construction and various characterizations of these processes. See Dawson (1993) for a recent survey. One way to define a Dawson–Watanabe process is as follows. Let E and A be as above, and let $\mathscr{M}_{\mathrm{f}}(E)$ denote the set of finite positive Borel measures on E with the topology of weak convergence. The Dawson–Watanabe process with motion operator A (or the A-DW process) is an $\mathscr{M}_{\mathrm{f}}(E)$-valued Markov process with generator given by

$$(\mathscr{L}^{\mathrm{DW}}\varphi)(\mu) = \frac{1}{2}\int_E \mu(\mathrm{d}x)\frac{\delta^2\varphi(\mu)}{\delta\mu(x)^2} + \int_E \mu(\mathrm{d}x)A\Big(\frac{\delta\varphi(\mu)}{\delta\mu(\cdot)}\Big)(x), \tag{1.2}$$

with $\mathscr{D}(\mathscr{L}^{\mathrm{DW}})$ defined similarly to $\mathscr{D}(\mathscr{L}^{\mathrm{FV}})$, except that $C^2(\mathbf{R}^k)$ is replaced by $C_c^2(\mathbf{R}^k)$.

Give $\Omega = C_{\mathscr{P}(E)}[0,\infty)$ and $\Xi = C_{\mathscr{M}_{\mathrm{f}}(E)}[0,\infty)$ the topologies of uniform convergence on compact sets, and let $\mathscr{F}$ and $\mathscr{G}$ denote the respective Borel σ-fields. Let $\{X_t,\ t\geqslant 0\}$ and $\{Y_t,\ t\geqslant 0\}$ be the canonical coordinate processes on Ω and Ξ, respectively, and put $\mathscr{F}_t = \sigma\{X_s : 0\leqslant s\leqslant t\}$ and $\mathscr{G}_t = \sigma\{Y_s : 0\leqslant s\leqslant t\}$.

There is an alternative characterization of the two processes that also suggests a close relationship. The A-FV process starting at $\mu \in \mathscr{P}(E)$ is the unique solution $P_\mu \in \mathscr{P}(\Omega)$ of the following martingale problem: (a) $P_\mu\{X_0 = \mu\} = 1$, and (b) for each $f \in \mathscr{D}(A)$, $\langle X_t, f\rangle - \int_0^t \langle X_s, Af\rangle\,\mathrm{d}s$ is an $\{\mathscr{F}_t\}$-martingale on $(\Omega, \mathscr{F}, P_\mu)$ with quadratic variation $\int_0^t(\langle X_s, f^2\rangle - \langle X_s, f\rangle^2)\,\mathrm{d}s$. The A-DW process starting at $\mu \in \mathscr{M}_{\mathrm{f}}(E)$ is the unique solution $Q_\mu \in \mathscr{P}(\Xi)$ of the following martingale problem: (a) $Q_\mu\{Y_0 = \mu\} = 1$, and (b) for each $f \in \mathscr{D}(A)$, $\langle Y_t, f\rangle - \int_0^t\langle Y_s, Af\rangle\,\mathrm{d}s$ is a $\{\mathscr{G}_t\}$-martingale on $(\Xi, \mathscr{G}, Q_\mu)$ with quadratic variation $\int_0^t\langle Y_s, f^2\rangle\,\mathrm{d}s$.

Not only are the two processes characterized similarly, but they have a number of properties in common as well. For example, when $E = \mathbf{R}^d$ and $A = \frac{1}{2}\Delta_d$, which corresponds to d-dimensional Brownian motion, then both processes have jointly continuous (in time and space) densities if $d = 1$ (Konno and Shiga, 1988; Reimers, 1992) and are singular at time t for all $t > 0$ if $d \geqslant 2$ (Dawson and Hochberg, 1979, 1982; Perkins, 1988). Furthermore, if $d \geqslant 2$, the Hausdorff dimension of the supports at time t is 2 for all $t > 0$ (Dawson and Hochberg, 1979, 1982; Zähle, 1988).

A better understanding of the above similarities can be obtained from the work of several authors. Donnelly and Kurtz (1995a,b) showed how to construct A-FV and A-DW processes as empirical measures for certain interacting particle systems. Etheridge and March (1991) and Perkins (1991) proved that, roughly speaking, an A-DW process conditioned to have total mass 1 is an A-FV process. Konno and Shiga (1988) obtained an $\tilde{A}$-FV process from an A-DW process via a random time change and normalization; here $\tilde{A}$ differs from A, however.

Our goal in this paper is to provide a way to compare A-FV and A-DW processes (with the same A) that is helpful in proving detailed sample path properties. In the notation introduced above, we work with product moments of the form $\mathbf{E}^{P_\mu}[\langle X_{t_1}, f_1\rangle \cdots \langle X_{t_n}, f_n\rangle]$ and $\mathbf{E}^{Q_\mu}[\langle Y_{t_1}, f_1\rangle \cdots \langle Y_{t_n}, f_n\rangle]$ for $n \geqslant 1$, $0 < t_1 \leqslant \cdots \leqslant t_n$, $f_1, \ldots, f_n \in B(E)$ (the space of bounded Borel functions on E), and appropriately chosen μ. When proving sample path properties, it is often the case that one needs good bounds on a number of these moments; in some cases, all of the moments are needed.

For Dawson–Watanabe processes, Dynkin (1988) has systematized the above moments in terms of directed binary graphs. The formulas one gets are quite cumbersome (especially when n is not small), but there are techniques for dealing with them. In the Fleming–Viot case, we will use the dual process to obtain a similar (but slightly more complicated) systematization of the moments in terms of certain "coalescent diagrams." The main theorem on comparing moments is the following.

Theorem 1.1. *Let $\{P_\mu : \mu \in \mathscr{P}(E)\}$ and $\{Q_\mu : \mu \in \mathscr{M}_f(E)\}$ correspond to the A-FV and A-DW processes as described above. Suppose $n \geqslant 1$, $0 < t_1 \leqslant \cdots \leqslant t_n$, and $f_1, \ldots, f_n \in B(E)$ are nonnegative.*

(a) *If $\mu \in \mathscr{P}(E)$, then*

$$\mathbf{E}^{P_\mu}[\langle X_{t_1}, f_1\rangle \cdots \langle X_{t_n}, f_n\rangle] \leqslant \mathbf{E}^{Q_\mu}[\langle Y_{t_1}, f_1\rangle \cdots \langle Y_{t_n}, f_n\rangle]. \tag{1.3}$$

(b) *If $\mu \in \mathscr{M}_f(E)$ with $\mu(E) > 0$, then*

$$\mathbf{E}^{Q_\mu}[\langle Y_{t_1}, f_1\rangle \cdots \langle Y_{t_n}, f_n\rangle] \leqslant C_n(t_{n-1}, \mu)\mathbf{E}^{P_{\bar\mu}}[\langle X_{t_1}, f_1\rangle \cdots \langle X_{t_n}, f_n\rangle], \tag{1.4}$$

where $\bar\mu = \mu(E)^{-1}\mu \in \mathscr{P}(E)$ and $C_n(t, \mu) = 2^{n-1}\mu(E)(\mu(E)^{n-1} \vee 1)e^{\binom{n}{2}t}$.

As a consequence, we have that any result about the A-DW process that can be proved by bounding moments of the above types will automatically hold for the A-FV process, and vice versa.

The bound in part (a) is perhaps the most useful since moment bounds for DW processes are already known and, as noted above, FV moments are a bit more complicated. It is this bound that will allow us to prove existence and continuity properties of local time for certain FV processes.

The rest of the paper is organized as follows. In Section 2 we obtain a moment formula for FV processes using dual processes and coalescent diagrams. In Section 3 we describe Dynkin's moment formula for DW processes and prove Theorem 1.1. Section 4 contains the applications to local time of FV processes, and Section 5 provides the proofs of these results.

2. Moment formula for FV processes

For the remainder of this paper, the only superprocess generator we will deal with is $\mathscr{L}^{\mathrm{FV}}$. To simplify notation, we simply write $\mathscr{L}$ for this operator from now on.

In the notation of the preceding section, we are interested in product moments of the form

$$\mathbf{E}^{P_\mu}[\langle X_{t_1}, f_1\rangle \cdots \langle X_{t_n}, f_n\rangle], \tag{2.1}$$

where $n \geqslant 1$, $0 < t_1 \leqslant \cdots \leqslant t_n$, $f_1, \ldots, f_n \in B(E)$, and $\mu \in \mathscr{P}(E)$.

For this we recall the function-valued dual process introduced by Dawson and Hochberg (1982). For each $k \geqslant 2$ and $1 \leqslant i < j \leqslant k$, we define $\Phi_{ij}^{(k)} : B(E^k) \to B(E^{k-1})$ by letting $\Phi_{ij}^{(k)}h$ be the function obtained from h by replacing x_j by x_i and renumbering the variables $x_{j+1}, \ldots, x_k$:

$$(\Phi_{ij}^{(k)}h)(x_1, \ldots, x_{k-1}) = h(x_1, \ldots, x_{j-1}, x_i, x_j, \ldots, x_{k-1}). \tag{2.2}$$

For example, $(\Phi_{12}^{(3)}h)(x_1, x_2) = h(x_1, x_1, x_2)$. For each $k \geqslant 1$, let $A^{(k)}$ be the generator of the semigroup $\{T_k(t)\}$ on $B(E^k)$ given by

$$T_k(t)h(x_1, \ldots, x_k) = \int_E \cdots \int_E h(y_1, \ldots, y_k) P_t(x_1, \mathrm{d}y_1) \cdots P_t(x_k, \mathrm{d}y_k), \tag{2.3}$$

where $P_t(x, \mathrm{d}y)$ is the transition function for the mutation process with generator A.

Now, for each $k \geqslant 1$, $h \in B(E^k)$, and $\mu \in \mathscr{P}(E)$, we define $\varphi_h \in B(\mathscr{P}(E))$ and $\varphi_\mu \in B(B(E^k))$ by

$$\varphi_h(\mu) = \langle \mu^k, h\rangle = \varphi_\mu(h), \tag{2.4}$$

where μ^k denotes the k-fold product measure $\mu \times \cdots \times \mu$, and $\langle \mu^k, h\rangle = \int_{E^k} h\,\mathrm{d}\mu^k$. We observe that, for each $k \geqslant 1$ and $h \in \mathscr{D}(A^{(k)})$, φ_h belongs to the domain of the closure $\overline{\mathscr{L}}$ of $\mathscr{L}$ and

$$(\overline{\mathscr{L}}\varphi_h)(\mu) = \sum_{1 \leqslant i < j \leqslant k} (\langle \mu^{k-1}, \Phi_{ij}^{(k)}h\rangle - \langle \mu^k, h\rangle) + \langle \mu^k, A^{(k)}h\rangle. \tag{2.5}$$

The dual Markov process assumes values in

$$\mathscr{H} = \bigcup_{k=1}^{\infty} B(E^k) \tag{2.6}$$

and has generator $\mathscr{L}^\#$ satisfying

$$(\overline{\mathscr{L}}\varphi_h)(\mu) = (\mathscr{L}^\#\varphi_\mu)(h). \tag{2.7}$$

By (2.5), this is a process that jumps from $h \in B(E^k)$ to $\Phi_{ij}^{(k)}h \in B(E^{k-1})$ at rate 1 $(1 \leqslant i < j \leqslant k)$. Between jumps it moves deterministically from $h \in B(E^k)$ to $T_k(t)h \in B(E^k)$ in time t.

Let us be more explicit about this process. Let $\Lambda_2, \Lambda_3, \ldots$ be independent exponential random variables with $\mathbf{E}[\Lambda_k] = 1/\binom{k}{2}$, and put $\Lambda_1 \equiv \infty$. Let $U_2, U_3, \ldots$ be independent discrete random variables (and independent of $\Lambda_2, \Lambda_3, \ldots$) with $\mathbf{P}\{U_k = (i,j)\} = 1/\binom{k}{2}$,

$1 \leqslant i < j \leqslant k$, and define the sequence $\Gamma_2, \Gamma_3, \ldots$ of independent random operators by $\Gamma_k = \Phi_{U_k}^{(k)}$. Then, given $m \geqslant 1$ and $g \in B(E^m)$, we define the dual process $\{Z_g(t),\ t \geqslant 0\}$ starting at g by

$$Z_g(t) = T_k(t - \Lambda_{k+1} - \cdots - \Lambda_m)\Gamma_{k+1} \cdots \Gamma_{m-1} T_{m-1}(\Lambda_{m-1})\Gamma_m T_m(\Lambda_m) g$$
$$\text{if} \quad \Lambda_{k+1} + \cdots + \Lambda_m \leqslant t < \Lambda_k + \cdots + \Lambda_m, \quad 1 \leqslant k \leqslant m. \tag{2.8}$$

It is important to note that $\{Z_g(t),\ (t,g) \in [0,\infty) \times \mathscr{H}\}$ is defined on a single probability space.

The duality relationship is expressed by the following result of Dawson and Hochberg (1982); see Ethier and Kurtz (1987) for a proof.

Lemma 2.1. *For each $m \geqslant 1$, $g \in B(E^m)$, $t \geqslant 0$, and $\mu \in \mathscr{P}(E)$,*

$$\mathbf{E}^{P_\mu}[\langle X_t^m, g\rangle] = \mathbf{E}[\langle \mu^{N_g(t)}, Z_g(t)\rangle], \tag{2.9}$$

where $N_g(t) = k$ if $Z_g(t) \in B(E^k)$.

With this background, we turn to the evaluation of (2.1). Fix $n \geqslant 1$, $0 < t_1 \leqslant \cdots \leqslant t_n$, and $f_1, \ldots, f_n \in B(E)$. Let $\{Z_g^{(m)}(t),\ (t,g) \in [0,\infty) \times \mathscr{H}\}$ $(m = 1, \ldots, n)$ be independent copies of $\{Z_g(t),\ (t,g) \in [0,\infty) \times \mathscr{H}\}$; they are defined on some probability space $(\Omega^*, \mathscr{F}^*, P^*)$ in terms of $\{(\Lambda_k^{(m)}, U_k^{(m)}), k \geqslant 2\}$ $(m = 1, \ldots, n)$, which are independent copies of $\{(\Lambda_k, U_k),\ k \geqslant 2\}$. If $0 < t_1 < \cdots < t_n$, define the $\mathscr{H}$-valued process $\{Z(t),\ 0 \leqslant t < t_n\}$ recursively as follows:

$$Z(t) = Z_{f_n}^{(1)}(t) = T_1(t) f_n, \quad 0 \leqslant t < t_n - t_{n-1}, \tag{2.10}$$

and, if $Z(t)$ has been defined for $0 \leqslant t < t_n - t_{n-m}$, where $1 \leqslant m \leqslant n-1$, then

$$Z(t) = Z_{Z((t_n - t_{n-m})-) \times f_{n-m}}^{(m+1)}(t - (t_n - t_{n-m})), \quad t_n - t_{n-m} \leqslant t < t_n - t_{n-m-1}. \tag{2.11}$$

Here, if $g \in B(E^k)$ and $f \in B(E)$, then $g \times f$ is the function in $B(E^{k+1})$ given by $(g \times f)(x_1, \ldots, x_{k+1}) = g(x_1, \ldots, x_k) f(x_{k+1})$; also, $t_0 = 0$.

If some of the times t_i are equal, we modify the above definitions in the obvious way. Specifically, if $t_{n-k-1} < t_{n-k} = \cdots = t_n$, where $1 \leqslant k \leqslant n-1$, then (2.10) becomes

$$Z(t) = Z_{f_n \times \cdots \times f_{n-k}}^{(1)}(t) = T_{k+1}(t)(f_n \times \cdots \times f_{n-k}), \quad 0 \leqslant t < t_n - t_{n-k-1}, \tag{2.10$'$}$$

and if $t_{n-m-k-1} < t_{n-m-k} = \cdots = t_{n-m}$, where $1 \leqslant m \leqslant n-2$ and $1 \leqslant k \leqslant n-m-1$, then (2.11) becomes

$$Z(t) = Z_{Z((t_n - t_{n-m})-) \times f_{n-m} \times \cdots \times f_{n-m-k}}^{(m+1)}(t - (t_n - t_{n-m})),$$
$$t_n - t_{n-m} \leqslant t < t_n - t_{n-m-k-1}. \tag{2.11$'$}$$

Lemma 2.2. *For each $n \geqslant 1$, $0 < t_1 \leqslant \cdots \leqslant t_n$, $f_1, \ldots f_n \in B(E)$, and $\mu \in \mathscr{P}(E)$,*

$$\mathbf{E}^{P_\mu}[\langle X_{t_1}, f_1\rangle \cdots \langle X_{t_n}, f_n\rangle] = \mathbf{E}^{P^*}[\langle \mu^{N(t_n-)}, Z(t_n-)\rangle], \tag{2.12}$$

where $\{Z(t),\ 0 \leqslant t < t_n\}$ is defined in terms of n, $t_1, \ldots, t_n$, and $f_1, \ldots, f_n$ by (2.10) and (2.11), and $N(t) = k$ if $Z(t) \in B(E^k)$.

Proof. Fix n, $t_1,\ldots,t_n$, $f_1,\ldots,f_n$, and μ as in the statement of the lemma. We extend $\{X_t,\ t\geqslant 0\}$, $\{Z_g^{(m)}(t),\ (t,g)\in[0,\infty)\times\mathscr{H}\}$ $(1\leqslant m\leqslant n)$, and $\{Z(t),\ 0\leqslant t<t_n\}$ to the product space $(\Omega\times\Omega^*,\mathscr{F}\times\mathscr{F}^*,P_\mu\times P^*)$ in the obvious way, denoting the extended processes by $\{\tilde{X}_t,\ t\geqslant 0\}$, $\{\tilde{Z}_g^{(m)}(t),\ (t,g)\in[0,\infty)\times\mathscr{H}\}$ $(1\leqslant m\leqslant n)$, and $\{\tilde{Z}(t),\ 0\leqslant t<t_n\}$. To keep the notation from getting completely out of hand, we give the details of the proof only in the case in which the times t_i are distinct. The obvious changes will handle the more general case.

First,

$$
\begin{aligned}
&\mathbf{E}^{P_\mu}[\langle X_{t_1},f_1\rangle\cdots\langle X_{t_n},f_n\rangle]\\
&\quad=\mathbf{E}^{P_\mu}[\langle X_{t_1},f_1\rangle\cdots\langle X_{t_{n-1}},f_{n-1}\rangle\mathbf{E}^{P_{X_{t_{n-1}}}}[\langle X_{t_n-t_{n-1}},f_n\rangle]]\\
&\quad=\mathbf{E}^{P_\mu\times P^*}[\langle\tilde{X}_{t_1},f_1\rangle\cdots\langle\tilde{X}_{t_{n-1}},f_{n-1}\rangle\langle\tilde{X}_{t_{n-1}},\tilde{Z}((t_n-t_{n-1})-)\rangle]
\end{aligned}
\tag{2.13}
$$

by the Markov property, Lemma 2.1, and Eq. (2.10). Next, for $m=1,\ldots,n-1$,

$$
\begin{aligned}
&\mathbf{E}^{P_\mu\times P^*}[\langle\tilde{X}_{t_1},f_1\rangle\cdots\langle\tilde{X}_{t_{n-m}},f_{n-m}\rangle\langle\tilde{X}_{t_{n-m}}^{\tilde{N}((t_n-t_{n-m})-)},\tilde{Z}((t_n-t_{n-m})-)\rangle]\\
&\quad=\mathbf{E}^{P_\mu\times P^*}[\langle\tilde{X}_{t_1},f_1\rangle\cdots\langle\tilde{X}_{t_{n-m-1}},f_{n-m-1}\rangle\langle\tilde{X}_{t_{n-m}}^{\tilde{N}((t_n-t_{n-m})-)+1},\\
&\qquad\tilde{Z}((t_n-t_{n-m})-)\times f_{n-m}\rangle]\\
&\quad=\mathbf{E}^{P_\mu\times P^*}[\langle\tilde{X}_{t_1},f_1\rangle\cdots\langle\tilde{X}_{t_{n-m-1}},f_{n-m-1}\rangle\\
&\qquad\mathbf{E}^{P_{\tilde{X}_{t_{n-m-1}}}\times P^*}[\langle\tilde{X}_{t_{n-m}-t_{n-m-1}}^{\tilde{N}((t_n-t_{n-m})-)+1},\tilde{Z}((t_n-t_{n-m})-)\times f_{n-m}\rangle]]\\
&\quad=\mathbf{E}^{P_\mu\times P^*}[\langle\tilde{X}_{t_1},f_1\rangle\cdots\langle\tilde{X}_{t_{n-m-1}},f_{n-m-1}\rangle\\
&\qquad\langle\tilde{X}_{t_{n-m-1}}^{\tilde{N}^{(m+1)}_{\tilde{Z}((t_n-t_{n-m})-)\times f_{n-m}}(t_{n-m}-t_{n-m-1})},\tilde{Z}^{(m+1)}_{\tilde{Z}((t_n-t_{n-m})-)\times f_{n-m}}(t_{n-m}-t_{n-m-1})\rangle]\\
&\quad=\mathbf{E}^{P_\mu\times P^*}[\langle\tilde{X}_{t_1},f_1\rangle\cdots\langle\tilde{X}_{t_{n-m-1}},f_{n-m-1}\rangle\\
&\qquad\langle\tilde{X}_{t_{n-m-1}}^{\tilde{N}((t_n-t_{n-m-1})-)},\tilde{Z}((t_n-t_{n-m-1})-)\rangle]
\end{aligned}
\tag{2.14}
$$

by Fubini's theorem, the Markov property, Lemma 2.1, and (2.11). Finally, since $t_0=0$,

$$
\mathbf{E}^{P_\mu\times P^*}[\langle\tilde{X}_{t_0}^{\tilde{N}((t_n-t_0)-)},\tilde{Z}((t_n-t_0)-)\rangle]=\mathbf{E}^{P^*}[\langle\mu^{N(t_n-)},Z(t_n-)\rangle],
\tag{2.15}
$$

and the proof is complete.

The process $\{Z(t),\ 0\leqslant t<t_n\}$ of Lemma 2.2 has a very simple probabilistic structure, which is related to Kingman's (1982) n-coalescent. Let us recall the latter process. For $k=1,\ldots,n$, let $\pi(n,k)$ be the set of partitions β of $\{1,\ldots,n\}$ into k nonempty subsets $\beta_1,\ldots,\beta_k$, labeled so that $\min\beta_1<\cdots<\min\beta_k$. The *n-coalescent* is a pure-jump Markov process in $\cup_{k=1}^n\pi(n,k)$ that jumps from $(\beta_1,\ldots,\beta_k)$ to

$$
(\beta_1,\ldots,\beta_{i-1},\beta_i\cup\beta_j,\beta_{i+1},\ldots,\beta_{j-1},\beta_{j+1},\ldots,\beta_k)
\tag{2.16}
$$

at rate 1, $1 \leqslant i < j \leqslant k$. Jumps are called *coalescences*. (It is often required that the initial state be $(\{1\},\ldots,\{n\})$, but we do not impose this requirement.)

Fix $n \geqslant 1$ and suppose for the moment that $0 = t_0 < t_1 < \cdots < t_n$. Consider the following variant $\{\zeta(t),\ 0 \leqslant t < t_n\}$ of the n-coalescent. Its state space is

$$\bigcup_{m=1}^{n} \bigcup_{k=1}^{m} \pi(m,k). \tag{2.17}$$

For $m = 1,\ldots,n$, $\{\zeta(t),\ t_n - t_{n-m+1} \leqslant t < t_n - t_{n-m}\}$ coincides with the restriction of an m-coalescent (with a suitable initial state) to the time interval $[0, t_{n-m-1} - t_{n-m})$. $\zeta(0) = (\{1\})$ and, if $1 \leqslant m \leqslant n-1$ and $\zeta((t_n - t_{n-m})-) = (\beta_1,\ldots,\beta_l) \in \pi(m,l)$, then $\zeta(t_n - t_{n-m}) = (\beta_1,\ldots,\beta_l,\{m+1\}) \in \pi(m+1,l+1)$. Loosely speaking, we begin with a single individual at time 0, and a new "immigrant" arrives at each of the times $t_n - t_{n-1} < \cdots < t_n - t_1$; between these fixed times, the process behaves like an ordinary coalescent. When we allow some of the times t_i to be equal, there is a similar process except that several immigrants can arrive at the same time. More precisely, if $t_{n-k-1} < t_{n-k} = \cdots = t_n$, where $1 \leqslant k \leqslant n-1$, then $\zeta(0) = (\{1\},\ldots,\{k+1\})$. If $t_{n-m-k-1} < t_{n-m-k} = \cdots = t_{n-m}$, where $1 \leqslant m \leqslant n-2$ and $1 \leqslant k \leqslant n-m-1$, and if $\zeta((t_n - t_{n-m})-) = (\beta_1,\ldots,\beta_l) \in \pi(m,l)$, then $\zeta(t_n - t_{n-m}) = (\beta_1,\ldots,\beta_l,\{m+1\},\ldots,\{m+k+1\}) \in \pi(m+k+1,l+k+1)$.

There is a natural correspondence between the sample paths of the process $\{\zeta(t),\ 0 \leqslant t < t_n\}$ just defined and those of the process $\{Z(t),\ 0 \leqslant t < t_n\}$ of Lemma 2.2. The sample paths of $\{Z(t),\ 0 \leqslant t < t_n\}$ are completely determined by the values of the random variables $(\Lambda_k^{(m)}, U_k^{(m)})$ ($k \geqslant 2$, $m = 1,\ldots,n$) defined just above (2.10), and the same random variables can be used to construct $\{\zeta(t),\ 0 \leqslant t < t_n\}$. Indeed, $\{\Lambda_k^{(m)},\ k \geqslant 2\}$ represents a sequence of inter-coalescence times, and $\{U_k^{(m)},\ k \geqslant 2\}$ describes the corresponding sequence of coalescences.

An example may help to clarify this.

Example 2.3. Let $n = 4$ and fix $0 < t_1 < t_2 < t_3 < t_4$, $f_1,\ldots,f_4 \in B(E)$, and $\mu \in \mathscr{P}(E)$. Let Γ be the event that $\Lambda_2^{(2)} > t_3 - t_2$ and

$$0 < \Lambda_3^{(3)} < t_2 - t_1 < \Lambda_3^{(3)} + \Lambda_2^{(3)}, \qquad U_3^{(3)} = (2,3), \tag{2.18}$$

$$0 < \Lambda_3^{(4)} < \Lambda_3^{(4)} + \Lambda_2^{(4)} < t_1, \qquad U_3^{(4)} = (1,3). \tag{2.19}$$

Fig. 1 shows a typical sample path of $\{\zeta(t),\ 0 \leqslant t < t_4\}$ on Γ, and it is clear from the definitions that

$$\begin{aligned}\mathbf{E}^{P^*}[\langle \mu^{N(t_4-)}, Z(t_4-)\rangle;\ \Gamma] = e^{-(t_3-t_2)} &\int_{t_1}^{t_2} \mathrm{d}s_3\, 3e^{-3(t_2-s_3)}(1/3)e^{-(s_3-t_1)} \\ &\int_0^{t_1} \mathrm{d}s_2\, 3e^{-3(t_1-s_2)}(1/3) \int_0^{s_2} \mathrm{d}s_1\, e^{-(s_2-s_1)} \\ &\langle \mu, T(s_1)\{T(s_2-s_1)[T(t_4-s_2)f_4 \cdot T(t_1-s_2)f_1] \\ &\cdot T(s_3-s_1)[T(t_3-s_3)f_3 \cdot T(t_2-s_3)f_2]\}\rangle,\end{aligned} \tag{2.20}$$

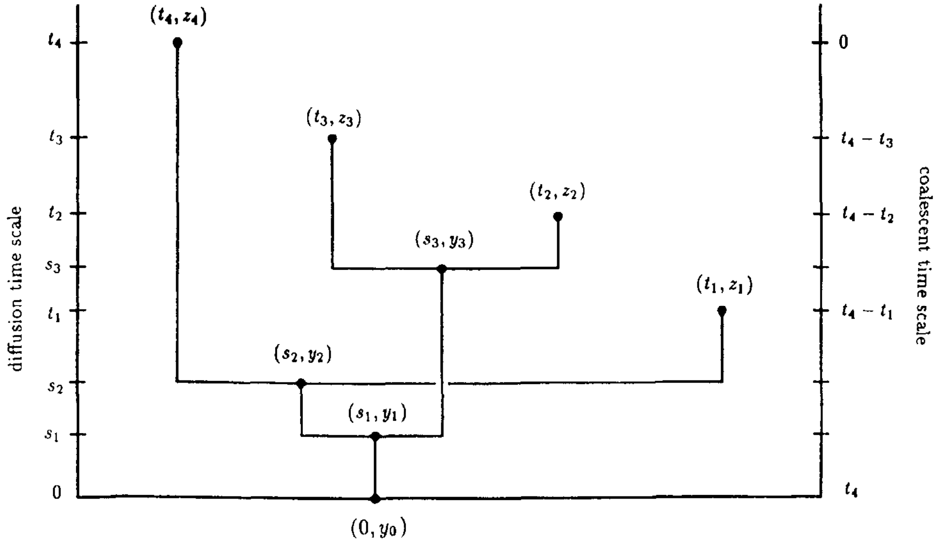

Fig. 1.

where $\{T(t)\} = \{T_1(t)\}$. Noting the labeling of the vertices of the graph in Fig. 1, we can rewrite the above quantity in terms of the transition function $P_t(x, \mathrm{d}y)$ as

$$\begin{aligned}&\int_{t_1}^{t_2} \mathrm{d}s_3 \int_0^{t_1} \mathrm{d}s_2 \int_0^{s_2} \mathrm{d}s_1 \int_E \mu(\mathrm{d}y_0)\\ &\times e^{-(t_3-t_2)}\, e^{-3(t_2-s_3)}\, e^{-(s_3-t_1)}\, e^{-3(t_1-s_2)}\, e^{-(s_2-s_1)}\\ &\times \int_E P_{s_1}(y_0, \mathrm{d}y_1) \int_E P_{s_2-s_1}(y_1, \mathrm{d}y_2) \int_E P_{s_3-s_1}(y_1, \mathrm{d}y_3)\\ &\times T(t_4 - s_2) f_4(y_2)\, T(t_3 - s_3) f_3(y_3)\, T(t_2 - s_3) f_2(y_3)\, T(t_1 - s_2) f_1(y_2). \qquad (2.21)\end{aligned}$$

We would like to generalize (2.20) and (2.21) and thereby express the moments (2.1) in a way that is particularly useful for comparing FV moments with DW moments. First, we need some additional terminology. Fix $n \geqslant 1$ and $0 < t_1 \leqslant \cdots \leqslant t_n$. Note that the sample paths of $\{\zeta(t),\ 0 \leqslant t < t_n\}$ have two kinds of jumps, namely, immigrations and coalescences. Let us say that two sample paths are *equivalent* if they have the same jumps in the same order but perhaps at different times. We refer to an equivalence class as a *coalescent diagram*, and we denote by $\mathscr{D}_n$ the set of all such coalescent diagrams. If $l \geqslant 1$ and $m_1, \ldots, m_l \geqslant 1$ satisfy $m_1 + \cdots + m_l = n$ and $t_1 = \cdots = t_{m_1} < t_{m_1+1} = \cdots = t_{m_1+m_2} < \cdots < t_{m_1+\cdots+m_{l-1}+1} = \cdots = t_n$, then its cardinality is

$$|\mathscr{D}_n| = \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{k_1+m_2} \sum_{k_3=1}^{k_2+m_3} \cdots \sum_{k_l=1}^{k_{l-1}+m_l} \prod_{j=1}^{l} \prod_{i=k_j+1}^{k_{j-1}+m_j} \binom{i}{2}, \qquad (2.22)$$

where $k_0 = 0$ and empty products are 1. For example, if $0 < t_1 < \cdots < t_n$, then $|\mathscr{D}_n| = 1, 2, 9, 79$ for $n = 1, 2, 3, 4$; if $0 < t_1 = \cdots = t_n$, then $|\mathscr{D}_n| = 1, 2, 7, 43$ for $n = 1, 2, 3, 4$. Fig. 1 can be thought of as an example of a coalescent diagram.

A coalescent diagram will have n *exit* vertices, at times $t_1 \leqslant \cdots \leqslant t_n$ (in the diffusion time scale), corresponding to the initial individual and the $n-1$ immigrants. It will have k $(1 \leqslant k \leqslant n)$ *entrance* vertices at time 0, and it will have $n-k$ *internal* vertices, corresponding to coalescences. We denote the entrance and internal vertices by (s_i, y_i) $(i = 0, 1, \ldots, n-1)$, where $s_0 = \cdots = s_{k-1} = 0$ and $0 < s_k < \cdots < s_{n-1}$. See Fig. 1, which has one entrance vertex labeled $(0, y_0)$, three internal vertices labeled $(s_1, y_1), (s_2, y_2), (s_3, y_3)$, and four exit vertices labeled $(t_1, z_1), \ldots, (t_4, z_4)$.

For each exit vertex (t_i, z_i) $(i = 1, \ldots, n)$, let $(s_{\beta(i)}, y_{\beta(i)})$ denote the starting vertex of the branch that ends at (t_i, z_i), and for each internal vertex (s_i, y_i) $(i = k, \ldots, n-1)$, let $(s_{\alpha(i)}, y_{\alpha(i)})$ denote the starting vertex of the branch that ends at (s_i, y_i).

The following moment formula for FV processes is a consequence of Lemma 2.2 and the above discussion; cf. Example 2.3.

Theorem 2.4. *For $n \geqslant 1$, $0 < t_1 \leqslant \cdots \leqslant t_n$, $f_1, \ldots, f_n \in B(E)$, and $\mu \in \mathscr{P}(E)$,*

$$\mathbf{E}^{P_\mu}[\langle X_{t_1}, f_1 \rangle \cdots \langle X_{t_n}, f_n \rangle] = \sum_{D \in \mathscr{D}_n} I_D, \tag{2.23}$$

where

$$\begin{aligned} I_D &= I_D(\mu) = I_D(\mu, t_1, \ldots, t_n, f_1, \ldots, f_n) \\ &= \prod_{i=k}^{n-1} \int \mathrm{d}s_i \prod_{i=0}^{k-1} \int_E \mu(\mathrm{d}y_i) \exp\left\{ -\int_0^{t_n} \binom{N(u)}{2} \mathrm{d}u \right\} \\ &\quad \prod_{i=k}^{n-1} \int_E P_{s_i - s_{\alpha(i)}}(y_{\alpha(i)}, \mathrm{d}y_i) \prod_{i=1}^{n} T(t_i - s_{\beta(i)}) f_i(y_{\beta(i)}), \end{aligned} \tag{2.24}$$

k being the number of entrance vertices of D; each $\mathrm{d}s_i$-integral ranges over the appropriate interval $(0, t_1)$ or (t_j, t_l), where $t_j < t_l$ are consecutive immigration times that are distinct, and, in addition, $s_k < \cdots < s_{n-1}$; furthermore, $N(u) = N_{D, t_1, \ldots, t_n, s_k, \ldots, s_{n-1}}(u)$ is the number of branches in D at time u in the coalescent time scale, when the exit and internal vertices are labeled in terms of $t_1, \ldots, t_n$ and $s_k, \ldots, s_{n-1}$.

3. Moment formula for DW processes, and proof of Theorem 1.1

We now describe Dynkin's (1988) moment formula for DW processes. Let $\mathscr{G}_n$ be the set of directed binary graphs with n exit vertices marked $1, 2, \ldots, n$ (sometimes referred to as *n-graphs*). Such a graph G is comprised of a set A of arrows and a set V of vertices. Each arrow begins at one vertex and ends at another. For each vertex v, let $a_+(v)$ denote the number of arrows that end at v, and $a_-(v)$ the number of arrows that begin at v. Thus, for each $v \in V$, there are exactly three possibilities: (i) $a_+(v) = 0$, $a_-(v) = 1$, (ii) $a_+(v) = 1$, $a_-(v) = 0$, or (iii) $a_+(v) = 1$, $a_-(v) = 2$. In case (i) we call v an *entrance* vertex, in case (ii) we call it an *exit* vertex, and in case (iii) we call it an *internal* vertex.

Permuting the labels of the exit vertices within a connected component yields a different n-graph, but the components themselves are unordered. For example, if we are given a 4-graph with two connected components having exits labeled 1,2 and 3,4, then any permutation of (1,2) and/or (3,4) gives a different 4-graph; however, switching (1,2) and (3,4) does not. So to every unlabeled directed binary graph with n exit vertices there corresponds no more than $n!$ (labeled) n-graphs. It is difficult to find a formula for $|\mathscr{G}_n|$ analogous to (2.22), but $|\mathscr{G}_n| = 1, 3, 19, 193$ for $n = 1, 2, 3, 4$.

For the purpose of stating Dynkin's theorem, we label the vertices of an n-graph G by two variables, one temporal and one spatial, in the same way we did for coalescent diagrams. Specifically, the n exit vertices are labeled by $(t_1, z_1), \ldots, (t_n, z_n)$ instead of by $1, \ldots, n$. If there are k $(1 \leqslant k \leqslant n)$ entrance vertices, there will be $n - k$ internal vertices. We denote the entrance and internal vertices by (s_i, y_i) $(i = 0, 1, \ldots, n-1)$, where $s_0 = \cdots = s_{k-1} = 0$. (Here we do not assume $s_k < \cdots < s_{n-1}$, as we did in Theorem 2.4.)

For each exit vertex (t_i, z_i) $(i = 1, \ldots, n)$, let $(s_{\beta(i)}, y_{\beta(i)})$ denote the starting vertex of the arrow that ends at (t_i, z_i), and for each internal vertex (s_i, y_i) $(i = k, \ldots, n-1)$, let $(s_{\alpha(i)}, y_{\alpha(i)})$ denote the starting vertex of the arrow that ends at (s_i, y_i).

We are now ready to state Dynkin's result for DW processes. Recall the notation $\{Y_t,\ t \geqslant 0\}$ and $\{Q_\mu : \mu \in \mathscr{M}_{\mathrm{f}}(E)\}$ introduced in Section 1.

Theorem 3.1. *For $n \geqslant 1$, $0 < t_1 \leqslant \cdots \leqslant t_n$, $f_1, \ldots, f_n \in B(E)$, and $\mu \in \mathscr{M}_{\mathrm{f}}(E)$,*

$$\mathbf{E}^{Q_\mu}[\langle Y_{t_1}, f_1\rangle \cdots \langle Y_{t_n}, f_n\rangle] = \sum_{G \in \mathscr{G}_n} J_G, \tag{3.1}$$

where

$$\begin{aligned} J_G &= J_G(\mu) = J_G(\mu, t_1, \ldots, t_n, f_1, \ldots, f_n) \\ &= \prod_{i=k}^{n-1} \int \mathrm{d}s_i \prod_{i=0}^{k-1} \int_E \mu(\mathrm{d}y_i) \\ &\quad \times \prod_{i=k}^{n-1} \int_E P_{s_i - s_{\alpha(i)}}(y_{\alpha(i)}, \mathrm{d}y_i) \prod_{i=1}^{n} T(t_i - s_{\beta(i)}) f_i(y_{\beta(i)}), \end{aligned} \tag{3.2}$$

k being the number of entrance vertices of G; the $\mathrm{d}s_i$-integrals are restricted only by the inequalities $s_{\alpha(i)} < s_i$ for $i = k, \ldots, n-1$ and $s_{\beta(i)} < t_i$ for $i = 1, \ldots, n$.

For example, if G is made up of several connected components, the ordering of the times s_i on one component is not affected by the times s_j on another.

Proof of Theorem 1.1. Define an equivalence relation between coalescent diagrams as follows. Two such diagrams are *equivalent* if they are graph-theoretically equivalent, that is, if they have the same tree structure (taking labeling of exit vertices into account), but $s_k, \ldots, s_{n-1} \in (0, t_n)$ are unrestricted. The corresponding equivalence classes are in one-to-one correspondence with the set of all *genealogical n-trees.* Such a tree can be represented in an obvious way by a (partially) nested collection of unordered pairs. For example, the coalescent diagram in Fig. 1 is one of three equivalent coalescent diagrams corresponding to the genealogical 4-tree given by $\{\{1,4\},\{2,3\}\}$; the other diagrams have 2 and 3 coalescing before 1 and 4 but after 1 arrives, or 1 and 4 coalescing before 2 and 3.

The collection of n-graphs can be partitioned in the same way, and equivalent n-graphs give the same integral J_G. If there are m coalescences ($0 \leqslant m \leqslant n-1$) in such an equivalence class (corresponding to a given genealogical n-tree), then there are 2^m elements in the equivalence class. So, in general, there will be at most 2^{n-1} equivalent n-graphs corresponding to a given genealogical n-tree.

To prove (a), note that since the exponential in (2.24) is bounded above by 1, we have $I_D \leqslant \tilde{I}_D$, where $\tilde{I}_D$ is obtained from I_D by removing this exponential factor. Now write $D \sim G$ when a coalescent diagram $D \in \mathscr{D}_n$ is equivalent to $G \in \mathscr{G}_n$ (i.e., when they correspond to the same genealogical n-tree). Then, using (2.24) and (3.2), $\sum_{D:D\sim G} \tilde{I}_D = J_G$, and (1.3) follows by summing over genealogical n-trees.

To prove (b), use the fact that $N(u) \leqslant n$ for all u and $N(u)=1$ if $t_{n-1} < u \leqslant t_n$ to obtain $\exp\{-\int_0^{t_n} \binom{N(u)}{2}\,\mathrm{d}u\} \geqslant e^{-\binom{n}{2}t_{n-1}}$. So, given $G \in \mathscr{G}_n$ and $\mu \in \mathscr{M}_{\mathrm{f}}(E)$,

$$\begin{aligned}
J_G(\mu) &= \prod_{i=k}^{n-1} \int \mathrm{d}s_i \prod_{i=0}^{k-1} \int_E \mu(\mathrm{d}y_i) \\
&\quad \prod_{i=k}^{n-1} \int_E P_{s_i - s_{\alpha(i)}}(y_{\alpha(i)}, \mathrm{d}y_i) \prod_{i=1}^{n} T(t_i - s_{\beta(i)}) f_i(y_{\beta(i)}) \\
&\leqslant \sum_{D:D\sim G} \mu(E)^k e^{\binom{n}{2}t_{n-1}} \prod_{i=k}^{n-1} \int \mathrm{d}s_i \prod_{i=0}^{k-1} \int_E \bar{\mu}(\mathrm{d}y_i) \exp\left\{-\int_0^{t_n} \binom{N(u)}{2}\,\mathrm{d}u\right\} \\
&\quad \prod_{i=k}^{n-1} \int_E P_{s_i - s_{\alpha(i)}}(y_{\alpha(i)}, \mathrm{d}y_i) \prod_{i=1}^{n} T(t_i - s_{\beta(i)}) f_i(y_{\beta(i)}) \\
&= \sum_{D:D\sim G} \mu(E)^k e^{\binom{n}{2}t_{n-1}} I_D(\bar{\mu}),
\end{aligned} \tag{3.3}$$

where we recall that k is the number of entrance vertices in G. The first product of $\mathrm{d}s_i$-integrals, after the first equality, are over the regions determined by G, whereas the second product of $\mathrm{d}s_i$-integrals, after the inequality, are over the regions determined by the choice of D. It now follows, by summing over $G \in \mathscr{G}_n$, that (1.4) holds. This completes the proof.

4. Applications

In recent years, superprocess local time has been investigated by a number of authors. See, for example, Iscoe (1986), Dynkin (1988), Sugitani (1989), Adler and Lewin (1992), and Krone (1993). In all of these works the superprocesses under consideration have been Dawson–Watanabe processes. It is known, for example, that local time exists in dimensions $d \leqslant 3$ for certain superdiffusions, including super-Brownian motion (i.e., the DW process with $E = \mathbf{R}^d$ and $A = \frac{1}{2}\Delta_d$). In the super-stable case (when the underlying motion is a symmetric stable process in $\mathbf{R}^d$ with index $\alpha \in (0,2]$), local time exists if $d < 2\alpha$. Other properties, such as joint continuity, Hölder continuity, Tanaka-type representations, and implications about sample path behavior of the superdiffusions can be found in the above references.

Our goal here is to establish existence and continuity properties of local time for the Fleming–Viot process with mutation process given by a diffusion in $\mathbf{R}^d$ with smooth uniformly elliptic coefficients and $d \leqslant 3$.

As in Section 1, we let X be the canonical coordinate process on $\Omega = C_{\mathscr{P}(E)}[0,\infty)$ and let $\mathscr{F} = \sigma\{X_t : t \geqslant 0\}$. The corresponding weighted occupation-time process is given by

$$\int_0^t X_s \, \mathrm{d}s. \tag{4.1}$$

Let P be a probability measure on $(\Omega, \mathscr{F})$. The *local time* of X on $(\Omega, \mathscr{F}, P)$, if it exists, is a process $\{L_t^x,\ (t,x) \in [0,\infty) \times \mathbf{R}^d\}$ with the property that, P-almost surely, $x \mapsto L_t^x$ is the density of (4.1) with respect to Lebesgue measure on $\mathbf{R}^d$, i.e.,

$$\int_0^t \langle X_s, f \rangle \, \mathrm{d}s = \int_{\mathbf{R}^d} f(y) L_t^y \, \mathrm{d}y \quad P\text{-a.s.}, \quad f \in B(\mathbf{R}^d), \tag{4.2}$$

for each $t \geqslant 0$. Informally, letting f be the Dirac delta "function" $\delta(\cdot - x)$, we have $L_t^x = \int_0^t \langle X_s, \delta(\cdot - x) \rangle \, \mathrm{d}s$. Thus, L_t^x can be thought of as the amount of mass that X puts on x during $[0,t]$. By standard results, one can choose a version of $\{L_t^x,\ (t,x) \in [0,\infty) \times \mathbf{R}^d\}$ such that $L_0^x = 0$ for all $x \in \mathbf{R}^d$, $t \mapsto L_t^x$ is nondecreasing and right continuous on $[0,\infty)$ for each $x \in \mathbf{R}^d$, and $x \mapsto L_t^x$ is Borel measurable on $\mathbf{R}^d$ for each $t \geqslant 0$. Existence of a local time means that the support of the measure-valued process "hits points" (cf. Krone, 1993).

We now consider local time for the A-FV process when A is the generator of a diffusion in $\mathbf{R}^d$ with smooth uniformly elliptic coefficients. Local time for the corresponding A-DW process has been studied in some of the references given at the beginning of this section.

We define the mutation operator A by

$$A = \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x) \frac{\partial}{\partial x_i} \tag{4.3}$$

acting on $\overline{C}^2(\mathbf{R}^d)$ (the space of bounded continuous functions on $\mathbf{R}^d$ with bounded continuous partial derivatives of orders 1 and 2), where the coefficients satisfy the following assumptions:

(a) the functions a_{ij}, b_i, $\partial a_{ij}/\partial x_j$, $\partial^2 a_{ij}/\partial x_i \partial x_j$, and $\partial b_i/\partial x_i$ are bounded and Hölder continuous on $\mathbf{R}^d$ $(i,j = 1,\ldots,d)$,

(b) there is a positive constant κ such that $\sum_{i,j=1}^{d} a_{ij}(x)\lambda_i\lambda_j \geqslant \kappa \sum_{i=1}^{d} \lambda_i^2$ for all real $\lambda_1, \ldots, \lambda_d$ and all $x \in \mathbf{R}^d$.

Let $p_t(x,y)$ denote the transition density (with respect to Lebesgue measure) for the diffusion with generator A (see Dynkin, 1965, Appendix). We will consider initial states for X in the set

$$\mathscr{P}_B(\mathbf{R}^d) = \left\{ \mu \in \mathscr{P}(\mathbf{R}^d) : \sup_{t>0,\, y\in\mathbf{R}^d} \int_{\mathbf{R}^d} p_t(x,y)\,\mu(\mathrm{d}x) < \infty \right\}. \tag{4.4}$$

It can be shown that $\mathscr{P}_B(\mathbf{R}^d)$ contains all $\mu \in \mathscr{P}(\mathbf{R}^d)$ with bounded densities with respect to Lebesgue measure. (See Krone (1993).)

Theorem 4.1. *Let $\{P_\mu : \mu \in \mathscr{P}(\mathbf{R}^d)\}$ correspond to the A-FV process as in Section 1 with $E = \mathbf{R}^d$ $(d \leqslant 3)$ and A given by* (4.3) *with coefficients satisfying* (a) *and* (b) *above. Suppose $\mu \in \mathscr{P}_B(\mathbf{R}^d)$. Then X on $(\Omega, \mathscr{F}, P_\mu)$ has a local time $\{L_t^x,\ (t,x) \in [0,\infty) \times \mathbf{R}^d\}$ that is jointly continuous P_μ-a.s. Moreover, the local time satisfies the following local Hölder condition for each index $\gamma \in (0, (2-d/2) \wedge 1)$. For every compact set $K \subset [0,\infty) \times \mathbf{R}^d$, there exist a positive random variable δ and a positive constant C such that, P_μ-a.s.,*

$$|L_t^x - L_s^y| \leqslant C|(t,x) - (s,y)|^\gamma \tag{4.5}$$

for all $(t,x), (s,y) \in K$ satisfying $|(t,x) - (s,y)| < \delta$, with $|\cdot|$ on the right side of (4.5) *denoting the $(d+1)$-dimensional Euclidean norm.*

Given the existence of a jointly continuous local time for X on $(\Omega, \mathscr{F}, P_\mu)$, some basic results follow as in Krone (1993). For example, let $S(\nu)$ be the closed support of the measure $\nu \in \mathscr{P}(E)$ and define the *x-level set* of X, for $x \in \mathbf{R}^d$, by

$$M_x = \{t \in [0,\infty) : x \in S(X_t)\}. \tag{4.6}$$

Then, P_μ-a.s., for a.e. $x \in \mathbf{R}^d$, $t \mapsto L_t^x$ increases only on M_x. Also, for fixed $x \in \mathbf{R}^d$, we have P_μ-a.s. that $L_t^x > 0$ implies $M_x \cap [0,t]$ is uncountable. For $T > 0$, define the *closed range* of X over $[0,T]$ to be

$$\overline{R}(T) = \overline{\cup_{t\in[0,T]} S(X_t)}. \tag{4.7}$$

Then, P_μ-a.s., $L_T^x = 0$ for all $x \in \overline{R}(T)^c$.

We also prove the following result, which makes rigorous the informal statement that $L_t^x = \int_0^t \langle X_s, \delta(\cdot - x)\rangle\,\mathrm{d}s$.

Theorem 4.2. *Under the assumptions of Theorem 4.1,*

$$L_t^x = \lim_{\varepsilon\to 0+} \int_0^t \langle X_s, p_\varepsilon(\cdot,x)\rangle\,\mathrm{d}s, \tag{4.8}$$

with the limit holding P_μ-a.s. and in $L^p(P_\mu)$ for each $t \geqslant 0$, $x \in \mathbf{R}^d$, and $1 \leqslant p < \infty$.

By results of Konno and Shiga (1988), when $d = 1$ we know that X on $(\Omega, \mathscr{F}, P_\mu)$ has a jointly continuous density $\xi(t,x)$, i.e.,

$$\langle X_t, f\rangle = \int_{\mathbf{R}} f(x)\xi(t,x)\,\mathrm{d}x \quad P_\mu\text{-a.s.}, \quad t \geqslant 0,\ f \in B(\mathbf{R}). \tag{4.9}$$

In this case, the local time is given directly by

$$L_t^x = \int_0^t \xi(s,x)\,\mathrm{d}s. \tag{4.10}$$

The proofs are given in the next section.

5. Remaining proofs

Here we prove the results stated in Section 4. We will give the proofs only for the case in which $A = \frac{1}{2}\Delta_d$. Using the estimates from Section 3 of Krone (1993), the same proofs can be made to work when A has the form (4.3). We will frequently use the following bounds for the d-dimensional Brownian transition density. We use the letter c to denote a generic constant (depending on d) whose value may change from line to line. If c depends on quantities other than d, that dependence will be indicated by means of a subscript.

Lemma 5.1. *For all $t,h > 0$ and $x,y,z \in \mathbf{R}^d$,*

(a) $|p_{t+h}(x,y) - p_t(x,y)| \leqslant cht^{-(d+2)/2}$,

(b) $|p_{t+h}(x,y) - p_t(x,y)| \leqslant ct^{-d/2}$,

(c) $\|p_t(\cdot,y)\|_2 \leqslant ct^{-d/4}$, *where $\|\cdot\|_2$ denotes the L^2 norm w.r.t. Lebesgue measure,*

(d) $|p_t(x,y) - p_t(x,z)| \leqslant c_\alpha|y-z|^\alpha t^{-\alpha/2}[p_{2t}(x,y) + p_{2t}(x,z)]$, *where $\alpha \in (0,1]$ is arbitrary.*

Proof. Both (a) and (b) follow from the inequality $|(\partial/\partial t)p_t(x,y)| \leqslant ct^{-(d+2)/2}$, which can be found in the Appendix to Dynkin (1965) or is easily verified directly. In (a), use the mean value theorem, and in (b), write the increment as an integral. Part (c) follows directly from the fact that, for any $b > 0$, $\int_{\mathbf{R}^d} \exp(-b|x|^2)\,\mathrm{d}x = (\pi/b)^{d/2}$. Part (d) can be found in (3.44) of Sugitani (1989).

Proof of Theorem 4.1. Fix $\mu \in \mathscr{P}_B(\mathbf{R}^d)$. Define, for $\varepsilon > 0$, $t > 0$, and $a \in \mathbf{R}^d$, the approximate local time

$$L_t^{a,\varepsilon} = \int_0^t \langle X_s, p_\varepsilon(\cdot,a)\rangle\,\mathrm{d}s. \tag{5.1}$$

To prove the existence of local time on $(\Omega, \mathscr{F}, P_\mu)$ we will show that

$$\mathbf{E}^{P_\mu}[(L_t^{a,\varepsilon} - L_t^{a,\varepsilon'})^n] \leqslant c_{n,\gamma,\mu,T}|\varepsilon - \varepsilon'|^{n\gamma} \tag{5.2}$$

for all $\varepsilon, \varepsilon' > 0$, $0 < t \leqslant T$, $a \in \mathbf{R}^d$, and $n = 2,4,6,\ldots$, where $d \leqslant 3$ and $0 < \gamma < (2 - d/2) \wedge 1$ is arbitrary. This immediately yields the existence of

$$L_t^{a,0} \equiv \lim_{\varepsilon \to 0+} L_t^{a,\varepsilon}, \tag{5.3}$$

with the limit holding in $L^p(P_\mu)$ for $1 \leqslant p < \infty$. To get a Hölder-continuous version, we will show that

$$\mathbf{E}^{P_\mu}[(L_t^{a,0} - L_t^{b,0})^n] \leqslant c_{n,\gamma,\mu,T}|a-b|^{n\gamma}, \tag{5.4}$$

$$\mathbf{E}^{P_\mu}[(L_t^{a,0} - L_s^{a,0})^n] \leqslant c_{n,\gamma,\mu,T}|t-s|^{n\gamma}, \tag{5.5}$$

for all $s,t \in [0,T]$ with $s < t$, $a,b \in \mathbf{R}^d$, and $n = 2,4,6,\ldots$, where $d \leqslant 3$ and $0 < \gamma < (2-d/2) \wedge 1$ is arbitrary.

The multiparameter version of the Kolmogorov–Čentsov theorem (see Karatzas and Shreve, 1988, Problem 2.2.9) implies that there is a jointly continuous modification of $L_t^{a,0}$, which we denote by L_t^a, and that this modification also obeys the required Hölder condition. Moreover, since the limit (5.3) holds uniformly for (t,a) in compact subsets of $[0,\infty) \times \mathbf{R}^d$, we conclude that $L_t^{a,0}$, and hence L_t^a, satisfies the occupation density property (4.2) (with P_μ in place of P).

If we were interested only in the limit (5.3) holding in $L^2(P_\mu)$, it would be enough to apply Theorem 1.1 to the proof of Proposition 3.2 in Sugitani (1989). In particular, the existence of local time for the A-FV process is easy. However, we are interested in obtaining convergence in $L^p(P_\mu)$ for all $p \geqslant 1$, because this will also allow us to use Theorem 1.1 to give a quick proof of (5.5). The proofs we give of (5.2) and (5.4) would work just as well for the A-DW process because of Theorem 1.1. The point is that hard work needs to be done for only one type of process. Unfortunately, the existing proofs for the DW case do not take this approach, so we must do some of the hard work here.

We will need to bound moments of the form

$$\mathbf{E}^{P_\mu}\left[\left(\int_r^u \langle X_v, f\rangle \,\mathrm{d}v\right)^n\right] \tag{5.6}$$

for certain choices of $f \in B(\mathbf{R}^d)$ and $0 \leqslant r < u$. Now

$$\begin{aligned}
&\mathbf{E}^{P_\mu}\left[\left(\int_r^u \langle X_v, f\rangle \,\mathrm{d}v\right)^n\right] \\
&= \int_r^u \cdots \int_r^u \mathbf{E}^{P_\mu}[\langle X_{t_1}, f\rangle \cdots \langle X_{t_n}, f\rangle]\,\mathrm{d}t_1 \cdots \mathrm{d}t_n \\
&= n! \int_r^u \mathrm{d}t_n \int_r^{t_n} \mathrm{d}t_{n-1} \cdots \int_r^{t_2} \mathrm{d}t_1\, \mathbf{E}^{P_\mu}[\langle X_{t_1}, f\rangle \cdots \langle X_{t_n}, f\rangle] \\
&= n! \sum_{D \in \mathscr{D}_n} \int_r^u \mathrm{d}t_n \int_r^{t_n} \mathrm{d}t_{n-1} \cdots \int_r^{t_2} \mathrm{d}t_1\, I_D(\mu, t_1, \ldots, t_n, f, \ldots, f),
\end{aligned} \tag{5.7}$$

where (cf. Theorem 2.4)

$$\begin{aligned}
I_D = {} & \prod_{i=k}^{n-1} \int \mathrm{d}s_i \prod_{i=0}^{k-1} \int_{\mathbf{R}^d} \mu(\mathrm{d}y_i) \prod_{i=k}^{n-1} \int_{\mathbf{R}^d} \mathrm{d}y_i \exp\left\{-\int_0^{t_n} \binom{N(u)}{2} \mathrm{d}u\right\} \\
& \times \prod_{i=k}^{n-1} p_{s_i - s_{\alpha(i)}}(y_{\alpha(i)}, y_i) \prod_{i=1}^{n} T(t_i - s_{\beta(i)}) f(y_{\beta(i)}).
\end{aligned} \tag{5.8}$$

To obtain a bound on (5.6), it is enough to consider I_D for each $D \in \mathscr{D}_n$. Since I_D factors into a product when D has several connected components, it is sufficient to deal with each connected coalescent diagram with n replaced by the number of exit vertices. With this simplification in mind, we will be treating only coalescent diagrams with a single entrance vertex $(0, y_0)$, and we denote the internal vertices by $(s_1, y_1), \dots, (s_{m-1}, y_{m-1})$, where m is the number of exit vertices.

To simplify the exposition of the proof, we first show how the computations are done in the case of Fig. 1. Once this is understood, the general case is easily handled.

We begin by proving that

$$\mathbf{E}^{P_\mu}[(L_t^{a,\varepsilon} - L_t^{b,\varepsilon})^n] \leqslant c_{n,\gamma,\mu,T} |a - b|^{n\gamma} \tag{5.9}$$

for all $\varepsilon > 0$, $0 < t \leqslant T$, $a, b \in \mathbf{R}^d$, and $n = 2, 4, 6, \dots$, where $d \leqslant 3$ and $0 < \gamma < (2 - d/2) \wedge 1$ is arbitrary. We do this by using (5.7) with $f(x) = p_\varepsilon(x, a) - p_\varepsilon(x, b)$, $r = 0$, and $u = t$. By Lemma 5.1(d),

$$\begin{aligned} |T(r)f(y)| &= |p_{r+\varepsilon}(y, a) - p_{r+\varepsilon}(y, b)| \\ &\leqslant c_\alpha |a - b|^\alpha r^{-\alpha/2} [p_{2(r+\varepsilon)}(y, a) + p_{2(r+\varepsilon)}(y, b)], \end{aligned} \tag{5.10}$$

where $\alpha \in (0, 1]$ is arbitrary.

Consider the case in which D is as in Fig. 1. By (2.21) we have

$$\begin{aligned} &\int_0^t \mathrm{d}t_4 \int_0^{t_4} \mathrm{d}t_3 \int_0^{t_3} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \, |I_D| \\ &\leqslant c_\alpha \, |a - b|^{4\alpha} \int_0^t \mathrm{d}t_4 \int_0^{t_4} \mathrm{d}t_3 \int_0^{t_3} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \\ &\times \int_{t_1}^{t_2} \mathrm{d}s_3 \int_0^{t_1} \mathrm{d}s_2 \int_0^{s_2} \mathrm{d}s_1 \int_{\mathbf{R}^d} \mu(\mathrm{d}y_0) \int_{\mathbf{R}^d} \mathrm{d}y_1 \int_{\mathbf{R}^d} \mathrm{d}y_2 \int_{\mathbf{R}^d} \mathrm{d}y_3 \\ &\times p_{s_1}(y_0, y_1) p_{s_2 - s_1}(y_1, y_2) p_{s_3 - s_1}(y_1, y_3) \\ &\times (t_4 - s_2)^{-\alpha/2} (t_3 - s_3)^{-\alpha/2} (t_2 - s_3)^{-\alpha/2} (t_1 - s_2)^{-\alpha/2} \\ &\times [p_{2(t_4 - s_2 + \varepsilon)}(y_2, a) + p_{2(t_4 - s_2 + \varepsilon)}(y_2, b)] \\ &\times [p_{2(t_3 - s_3 + \varepsilon)}(y_3, a) + p_{2(t_3 - s_3 + \varepsilon)}(y_3, b)] \\ &\times [p_{2(t_2 - s_3 + \varepsilon)}(y_3, a) + p_{2(t_2 - s_3 + \varepsilon)}(y_3, b)] \\ &\times [p_{2(t_1 - s_2 + \varepsilon)}(y_2, a) + p_{2(t_1 - s_2 + \varepsilon)}(y_2, b)]. \end{aligned} \tag{5.11}$$

Now integrate out y_0, using the fact that $\mu \in \mathscr{P}_B(\mathbf{R}^d)$. Only two factors involving y_1 will remain. Use Cauchy–Schwarz and Lemma 5.1(c) to get

$$\begin{aligned} \int_{\mathbf{R}^d} \mathrm{d}y_1 \, p_{s_2 - s_1}(y_1, y_2) \, p_{s_3 - s_1}(y_1, y_3) &\leqslant \|p_{s_2 - s_1}(\cdot, y_2)\|_2 \, \|p_{s_3 - s_1}(\cdot, y_3)\|_2 \\ &\leqslant c \, (s_2 - s_1)^{-d/4} (s_3 - s_1)^{-d/4}. \end{aligned} \tag{5.12}$$

In doing this, we have removed one of the factors involving y_2 (which corresponds to the next highest internal vertex). Applying Cauchy–Schwarz, Minkowski, and

Lemma 5.1(c) again, the remaining factors involving y_2 integrate as follows:

$$\int_{\mathbf{R}^d} \mathrm{d}y_2\,[p_{2(t_4-s_2+\varepsilon)}(y_2,a)+p_{2(t_4-s_2+\varepsilon)}(y_2,b)] \\ [p_{2(t_1-s_2+\varepsilon)}(y_2,a)+p_{2(t_1-s_2+\varepsilon)}(y_2,b)] \\ \leqslant c\,(t_4-s_2)^{-d/4}(t_1-s_2)^{-d/4}. \tag{5.13}$$

Repeat this procedure, integrating out y_3 as well. We are left with

$$\int_0^t \mathrm{d}t_4 \int_0^{t_4} \mathrm{d}t_3 \int_0^{t_3} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1\,|I_D| \\ \leqslant c_{\alpha,\mu}|a-b|^{4\alpha} \int_0^t \mathrm{d}t_4 \int_0^{t_4} \mathrm{d}t_3 \int_0^{t_3} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \int_{t_1}^{t_2} \mathrm{d}s_3 \int_0^{t_1} \mathrm{d}s_2 \int_0^{s_2} \mathrm{d}s_1 \\ \times(s_2-s_1)^{-d/4}(s_3-s_1)^{-d/4}(t_4-s_2)^{-d/4-\alpha/2} \\ \times(t_3-s_3)^{-d/4-\alpha/2}(t_2-s_3)^{-d/4-\alpha/2}(t_1-s_2)^{-d/4-\alpha/2}. \tag{5.14}$$

Rearranging the time integrals so that one starts at the top of the diagram and works down, it is easy to see that the remaining multiple integral is bounded by a function of t that is bounded on $[0,T]$ when $d\leqslant 3$ and $\alpha<(2-d/2)\wedge 1$. Thus,

$$\int_0^t \mathrm{d}t_4 \int_0^{t_4} \mathrm{d}t_3 \int_0^{t_3} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1\,|I_D| \leqslant c_{\alpha,\mu,T}|a-b|^{4\alpha}. \tag{5.15}$$

We now consider the more general case in which $D\in\mathscr{D}_m$ is connected. The result is trivial when $m=1$, so we assume $m\geqslant 2$. As before, we begin by bounding the $\mu(\mathrm{d}y_0)$-integral. Now, as far as the spatial variables are concerned, the factor corresponding to the lowest branch is gone. If (s_1,y_1) is the lowest internal vertex, then there are only two factors remaining that contain y_1; these correspond to the two branches going up from (s_1,y_1). There are three possibilities for this product of two factors, depending on how many of the branches end at exit vertices.

Bound the $\mathrm{d}y_1$-integral of this product, with the help of Lemma 5.1, as follows. First,

$$\int_{\mathbf{R}^d} \mathrm{d}y_1\, p_{s_2-s_1}(y_1,y_2)\,p_{s_3-s_1}(y_1,y_3) \leqslant \|p_{s_2-s_1}(\cdot,y_2)\|_2\|p_{s_3-s_1}(\cdot,y_3)\|_2 \\ \leqslant c\,(s_2-s_1)^{-d/4}(s_3-s_1)^{-d/4}; \tag{5.16}$$

second,

$$\int_{\mathbf{R}^d} \mathrm{d}y_1\, p_{s_2-s_1}(y_1,y_2)\,|T(t_i-s_1)f(y_1)| \\ \leqslant c_\alpha|a-b|^\alpha(t_i-s_1)^{-\alpha/2}\int_{\mathbf{R}^d}\mathrm{d}y_1\,p_{s_2-s_1}(y_1,y_2) \\ \times[p_{2(t_i-s_1+\varepsilon)}(y_1,a)+p_{2(t_i-s_1+\varepsilon)}(y_1,b)] \\ \leqslant c_\alpha|a-b|^\alpha(t_i-s_1)^{-\alpha/2}\|p_{s_2-s_1}(\cdot,y_2)\|_2(\|p_{2(t_i-s_1+\varepsilon)}(\cdot,a)\|_2+\|p_{2(t_i-s_1+\varepsilon)}(\cdot,b)\|_2) \\ \leqslant c_\alpha|a-b|^\alpha(t_i-s_1)^{-d/4-\alpha/2}(s_2-s_1)^{-d/4}; \tag{5.17}$$

third,

$$\begin{aligned}
&\int_{\mathbf{R}^d} \mathrm{d}y_1\, |T(t_i - s_1)f(y_1)|\, |T(t_j - s_1)f(y_1)| \\
&\quad \leqslant c_\alpha |a-b|^{2\alpha}(t_i - s_1)^{-\alpha/2}(t_j - s_1)^{-\alpha/2} \\
&\qquad \times \int_{\mathbf{R}^d} \mathrm{d}y_1\, [p_{2(t_i - s_1 + \varepsilon)}(y_1, a) + p_{2(t_i - s_1 + \varepsilon)}(y_1, b)] \\
&\qquad \times [p_{2(t_j - s_1 + \varepsilon)}(y_1, a) + p_{2(t_j - s_1 + \varepsilon)}(y_1, b)] \\
&\quad \leqslant c_\alpha |a-b|^{2\alpha}(t_i - s_1)^{-d/4-\alpha/2}(t_j - s_1)^{-d/4-\alpha/2}.
\end{aligned} \tag{5.18}$$

In each of the three cases, we get a bound with a factor $|a-b|^\alpha$ for each exit vertex, along with other factors involving only time variables. Now proceed up the diagram to the next internal vertex, say at (s_2, y_2). There are only two factors remaining that involve y_2. Bound the $\mathrm{d}y_2$-integral as above, and so on.

Since there are m exit vertices, after all the spatial variables have been integrated out, we will have a bound of the form

$$c_{\alpha,\mu,D}|a-b|^{m\alpha} \times \text{remaining time integrals.} \tag{5.19}$$

Each of the temporal factors is positive over the region of integration with exponent larger than -1. As in the example above, start at the top of the diagram and work down, integrating out the time variables one at a time. Each such integral can be bounded by $c_\alpha T$, and hence the next integrand contains only one factor involving the variable of integration. Conclude that, if $d \leqslant 3$ and $\alpha < (2 - d/2) \wedge 1$,

$$\int \mathrm{d}\mathbf{t}\, |I_D| \leqslant c_{\alpha,\mu,T,D}|a-b|^{m\alpha}, \tag{5.20}$$

generalizing (5.15). In view of (5.7), this implies (5.9).

We now prove (5.2). With D as in Figure 1, we apply (5.7) with $f(x) = p_\varepsilon(x,a) - p_{\varepsilon'}(x,a)$, $r = 0$, and $u = t$. Use (a) and (b) from Lemma 5.1 to get, for arbitrary $\delta \in (0,1)$,

$$\begin{aligned}
|T(r)f(y)| &= |p_{r+\varepsilon}(y,a) - p_{r+\varepsilon'}(y,a)|^\delta |p_{r+\varepsilon}(y,a) - p_{r+\varepsilon'}(y,a)|^{1-\delta} \\
&\leqslant c\, |\varepsilon - \varepsilon'|^\delta r^{-d/2-\delta}.
\end{aligned} \tag{5.21}$$

We apply this to the factors $T(t_4 - s_2)f(y_2)$ and $T(t_3 - s_3)f(y_3)$. In general, when an internal vertex branches into two exit vertices, we apply (5.21) to the factor corresponding to the higher exit vertex. Bound $|T(t_2 - s_3)f(y_3)|$ by $[p_{t_2 - s_3 + \varepsilon}(y_3, a) + p_{t_2 - s_3 + \varepsilon'}(y_3, a)]$, and give a similar bound for $|T(t_1 - s_2)f(y_2)|$. Now integrate out y_0 as before. Then proceed up the diagram, integrating out y_1, y_2, and y_3 as follows. If the diagram splits into two nonexiting branches at (s_i, y_i), bound the integral using Cauchy–Schwarz, as before. If the diagram splits into two exiting branches at (s_i, y_i), there will be only one factor involving y_i; it is of the form $[p_{t_j - s_i + \varepsilon}(y_i, a) + p_{t_j - s_i + \varepsilon'}(y_i, a)]$. The integral in this case equals 2. Thus, we have

$$\int \mathrm{d}t\,|I_D| \leqslant c\,|\varepsilon - \varepsilon'|^{2\delta} \int \mathrm{d}t \int \mathrm{d}s\,(t_4 - s_2)^{-d/2-\delta}(t_3 - s_3)^{-d/2-\delta}$$
$$(s_2 - s_1)^{-d/4}(s_3 - s_1)^{-d/4}, \tag{5.22}$$

and it follows easily that, if $d \leqslant 3$ and $0 < \delta < (2 - d/2) \wedge 1$,

$$\int \mathrm{d}t\,|I_D| \leqslant c_{\delta,\mu,T}|\varepsilon - \varepsilon'|^{2\delta}. \tag{5.23}$$

The general case can be handled as above. The only significant difference is that it is possible for the diagram to split into one exiting branch and one nonexiting branch at (s_i, y_i). In this case, use the bound (5.21) and notice that the remaining factor involving y_i integrates to 1.

Notice that (5.4) follows from (5.9) and (5.3).

It remains to prove (5.5). Let l_t^x be the local time for the A-DW process on $(\Xi, \mathscr{G}, Q_\mu)$. It is shown in Krone (1993) that

$$\int_0^t \langle Y_s, p_\varepsilon(\cdot, x)\rangle\,\mathrm{d}s \to l_t^x \quad \text{in } L^p(Q_\mu), \tag{5.24}$$

for all $p \geqslant 1$, and that

$$\mathbf{E}^{Q_\mu}[|l_t^x - l_s^x|^n] \leqslant c_{n,\gamma,\mu,T}|t - s|^{n\gamma} \tag{5.25}$$

for all $s, t \in [0, T]$, $x \in \mathbf{R}^d$, and $n = 2, 4, 6, \ldots$, where $d \leqslant 3$ and $0 < \gamma < (2 - d/2) \wedge 1$ is arbitrary. Therefore, for all $s, t \in [0, T]$ with $s < t$, $a \in \mathbf{R}^d$, and n, d, and γ as in (5.25),

$$\begin{aligned}
\mathbf{E}^{P_\mu}[|L_t^a - L_s^a|^n] &= \lim_{\varepsilon \to 0} \mathbf{E}^{P_\mu}\left[\left(\int_s^t \langle X_u, p_\varepsilon(\cdot, a)\rangle\,\mathrm{d}u\right)^n\right] \\
&= \lim_{\varepsilon \to 0} \int_s^t \cdots \int_s^t \mathbf{E}^{P_\mu}\left[\prod_{i=1}^n \langle X_{u_i}, p_\varepsilon(\cdot, a)\rangle\right] \mathbf{du} \\
&\leqslant \lim_{\varepsilon \to 0} \int_s^t \cdots \int_s^t \mathbf{E}^{Q_\mu}\left[\prod_{i=1}^n \langle Y_{u_i}, p_\varepsilon(\cdot, a)\rangle\right] \mathbf{du} \\
&= \mathbf{E}^{Q_\mu}[|l_t^a - l_s^a|^n] \\
&\leqslant c_{n,\gamma,\mu,T}|t - s|^{n\gamma},
\end{aligned} \tag{5.26}$$

where the first inequality uses (1.3). Note that we were able to apply Theorem 1.1 directly (without actually evaluating the integrals) because $f = p_\varepsilon(\cdot, a) \geqslant 0$, unlike the f's used previously. This completes the proof.

Proof of Theorem 4.2. The L^p convergence was established in the proof of Theorem 4.1 (see (5.3)). As for the a.s. convergence, note that (5.2) holds for all $\varepsilon, \varepsilon' \geqslant 0$, so we can apply the Kolmogorov–Čentsov theorem (Karatzas and Shreve, 1988, Theorem 2.2.8) to get the desired result.

References

R.J. Adler and M. Lewin, Local time and Tanaka formulae for super Brownian and super stable processes, Stochastic Process. Appl. 41 (1992) 45–67.

D.A. Dawson, Stochastic evolution equations and related measure processes, J. Multivariate Anal. 5 (1975) 1–52.

D.A. Dawson, Measure-valued Markov processes, in: D.A. Dawson, B. Maisonneuve and J. Spencer, eds., Ecole d'Eté de Probabilités de Saint-Flour XXI – 1991, Lecture Notes in Mathematics, Vol. 1541 (Springer, Berlin, 1993) pp. 1–260.

D.A. Dawson and K.J. Hochberg, The carrying dimension of a stochastic measure diffusion, Ann. Probab. 7 (1979) 693–703.

D.A. Dawson and K.J. Hochberg, Wandering random measures in the Fleming–Viot model, Ann. Probab. 10 (1982) 554–580.

P. Donnelly and T.G. Kurtz, A countable representation of the Fleming–Viot measure-valued diffusion (1995a), Preprint.

P. Donnelly and T.G. Kurtz, Particle representations for measure-valued population models (1995b), Preprint.

E.B. Dynkin, Markov Processes, Vol. II (Springer, Berlin, 1965).

E.B. Dynkin, Representation for functionals of superprocesses by multiple stochastic integrals, with applications to self-intersection local times, Astérisque 157–158 (1988) 147–171.

A. Etheridge and P. March, A note on superprocesses, Probab. Theory Related Fields 89 (1991) 141–147.

S.N. Ethier and T.G. Kurtz, The infinitely-many-alleles model with selection as a measure-valued diffusion, in: M. Kimura, G. Kallianpur and T. Hida, eds., Stochastic Models in Biology, Lecture Notes in Biomathematics, Vol. 70 (Springer, Berlin, 1987) pp. 72–86.

S.N. Ethier and T.G. Kurtz, Fleming–Viot processes in population genetics, SIAM J. Control Optim. 31 (1993) 345–386.

P.J. Fitzsimmons, Construction and regularity of measure-valued branching processes, Israel J. Math. 64 (1988) 337–361.

W.H. Fleming and M. Viot, Some measure-valued Markov processes in population genetics theory, Indiana Univ. Math. J. 28 (1979) 817–843.

I. Iscoe, Ergodic theory and a local occupation time for measure-valued critical branching Brownian motion, Stochastics 18 (1986) 197–243.

I. Karatzas and S.E. Shreve, Brownian Motion and Stochastic Calculus (Springer, New York, 1988).

J.F.C. Kingman, The coalescent, Stochastic Process. Appl. 13 (1982) 235–248.

N. Konno and T. Shiga, Stochastic partial differential equations for some measure-valued diffusions, Probab. Theory Related Fields 79 (1988) 201–225.

S.M. Krone, Local times for superdiffusions, Ann. Probab. 21 (1993) 1599–1623.

E.A. Perkins, A space-time property of a class of measure-valued branching diffusions, Trans. Amer. Math. Soc. 305 (1988) 743–795.

E.A. Perkins, Conditional Dawson–Watanabe processes and Fleming–Viot processes, in: E. Cinlar and M.J. Sharpe, eds., Seminar on Stochastic Processes 1991 (Birkhauser, Basel, 1992) pp. 143–155.

M. Reimers, A new result on the support of the Fleming–Viot process, proved by non-standard construction (1992), Unpublished.

S. Sugitani, Some properties for the measure-valued branching diffusion processes, J. Math. Soc. Japan 41 (1989) 437–462.

S. Watanabe, A limit theorem of branching processes and continuous state branching processes, J. Math. Kyoto Univ. 8 (1968) 141–167.

U. Zähle, Self-similar random measures, I. Notion, carrying Hausdorff dimension, and hyperbolic distribution, Probab. Theory Related Fields 80 (1988) 79–100.